Development and Multi-Institutional Implementation of Coding and Transmission Standards for Health Outcomes Data

Richard E. Ward, M.D., M.B.A. (1), Cynthia H. MacWilliam, M.B.A. (1)
Eric Ye, M.S. (1), Andrew N. Russman, M.A. (1)
Reuben R. Richards, M.A. (2), Michael Huber, M.H.S.A. (3)

(1) Center for Clinical Effectiveness, Henry Ford Health System, Detroit, Michigan
(2) American Medical Group Association, Alexandria, Virginia
(3) Health Outcomes Institute, Bloomington, Minnesota

Health care delivery institutions and providers, employers, and government agencies throughout the U.S. are sharing information and pooling resources in an attempt to produce reliable measurements of health outcomes. The rapid rate of growth in the collection and pooling of health outcomes data has prompted the need for standardization. The work of health care organizations and consortiums pooling data would be greatly facilitated by widely accepted standards for the coding and transmitting of outcomes data. Moreover, standards allow for the inter-operation of data capture products, data analysis tools, and data pooling services developed by a variety of different vendors. The Health Outcomes Institute (HOI) and Henry Ford Health System (HFHS) have developed and maintain a database of health outcomes questions which provides a mechanism for uniquely coding data elements. HFHS and the American Medical Group Association (AMGA) have created a software tool to facilitate the conversion and transmission of health outcomes data elements in an American Society for Testing and Materials (ASTM)/Health Level Seven (HL7) format, which incorporates HOI question standards.

INTRODUCTION

Growing Role of Outcomes Measurement

The systematic measurement of the outcomes of medical and surgical treatments in clinical trials has been at the heart of science-based medicine for more than a century. Clinical trialists, epidemiologists and health services researchers have developed numerous outcomes measurement instruments which have been shown to be valid and reliable research tools. Over the last decade, there has been an explosion of interest in extending the use of such instruments to routine medical practice. In his 1988 Shattuck Lecture [1], Paul Ellwood, M.D. coined the term "outcomes management" and articulated the need for "a new universal language to communicate hurting, functioning, working, interacting, and living" in order to manage health care services and advance the "good of the patient."

Since Ellwood's lecture, many health care delivery institutions, professional societies, employer coalitions, government agencies, and accreditation organizations have initiated outcomes measurement activities. Almost all of these diverse activities fall into four categories: effectiveness research, accountability, practice improvement, and patient care. Effectiveness research involves an evaluation of the outcomes of a clinical practice or a comparison of alternative practices as they perform under routine conditions [2,3]. Accountability-oriented outcomes measurement involves an evaluation of a health care practitioner or organization [4,5]. Practice improvement-oriented outcomes measurement involves measurement over time to assess improvement or measurement across institutions or practitioners to identify "best practices" or "benchmark sites" [6]. Patient care-oriented outcomes measurement involves the use of outcomes data collected from a specific patient to guide decision-making regarding the care of that patient [7,8].

Over the past five years, a new industry has come into existence to provide the goods and services to support outcomes measurement. Participants in this industry include system developers and vendors offering products to support the collection of structured data from patients and care providers using optical mark readers, image scanners, fax machines, touch screens, hand-held devices, and conventional PC user interfaces. Other participants offer products and services for study management, data pooling, data analysis, comparative data brokerage, cost-effectiveness or "pharmaco-economic" analysis, and prognostic modeling.

0195—4210/96/$5.00

The Need for Coding and Transmission Standards

The work of health care organizations and consortiums pooling data from numerous participating institutions would be greatly facilitated by widely accepted standards for the coding of outcomes data. Without such standards, data pooling centers must develop and maintain interfaces for every version of every data collection form for every health condition under study at every institution. Moreover, without standards, data capture products, data analysis products, data pooling services, and others involved in data aggregation cannot inter-operate.

The purpose of this paper is to describe the development and multi-instiutional implementation of coding and transmission standards to solve health outcomes data pooling, capture and analysis problems.

DESIGN OBJECTIVES AND CONSTRAINTS

We identified five key objectives for outcomes coding and transmission standards:

1. Assist in the process of instrument development;
2. Facilitate multi-institutional consensus on the use of common measurement instruments (or components of instruments) which are valid, and reliable;
3. Foster the widespread dissemination of public domain measurement instruments;
4. Advance ongoing improvement in measurement instruments without requiring the continuous restructuring of existing databases;
5. Enable the development of a competitive market for compatible outcomes management products and services;

We identified the following design constraints:

1. Not Static. Although standardizing outcomes measurement instruments themselves may be appealing for comparability across institutions, we recognized that for the foreseeable future, outcomes instruments would remain a moving target as numerous parties develop new instruments and make frequent modifications to existing instruments.
2. No Waiting. When developing or modifying instruments, people will be unwilling to wait before implementing design changes.
3. Not Unfamiliar. To promote adoption of the standards, we should not create a new, unfamiliar method for coding and transmitting data or a new, unfamiliar scheme for grouping semantically similar codes by subject.

DESIGN SOLUTION

Question-Based Approach.

Most existing outcomes data systems are based on a "forms-based" data model, usually representing outcomes data as a table with rows corresponding to completed surveys and columns corresponding to the "data elements" or "items" included in each survey instrument. If more than one outcomes instrument is used, then multiple tables are used to represent the various instruments. As instruments are modified over time, new tables are needed to represent each new version. Early attempts to develop standards for outcomes data transmission involved standardizing the field layout for each version of each instrument. The large number of instruments and the frequency of modifications made this approach unworkable.

An alternative approach could have been to assign a unique code to each response option for each data element, and treat the data from an outcomes survey as a stream of "answer codes", analogous to treatment of lab test results in an HL7 transmission. This approach could handle modification to existing instruments by adding new codes for new response options. However, a new response option for a subjective item on an outcomes survey instrument could create ambiguity to the meaning of pre-existing response options. For example, if an instrument contains a question "how do you feel today?" and the response options are "good" and "bad", and a new version of the instrument adds the response options "very good" and "excellent", then the meaning of a "good" response may be different on the new and old versions. Thus, for subjective information such as health status and satisfaction data from patients or clinicians, the visual and semantic context of the data entry must be recorded to avoid meaningful ambiguity or uncertainty of interpretation.

The *question-based approach* which we incorporated into the outcomes data coding and transmission standards, is intended to represent a reasonable, practical compromise. This approach assures that the actual words seen by the survey taker, the emphasis of the words, and the general layout (the relationships between the words, the response options and basic graphical components) are consistent for all data coded for a given form or question. Primary database structure, however,

exists at the question level and not form level, allowing for the flexible creation of forms.

Anatomy of a Question

Questions contain one or more *elements* which contain *response options* .

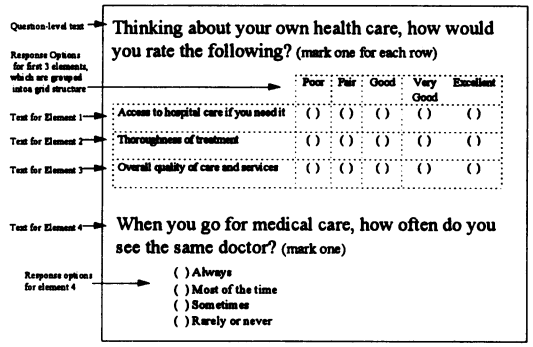

Figure 1. Anatomy of a question. [9]

For example (Figure 1), if a question asked: "Thinking about your own health care, how would you rate the following?", the first element might be "Access to hospital care if you need it" and the response options for the first element might be "Poor", "Fair", "Good", "Very Good", and "Excellent". These response options could have numeric values of 1, 2, 3, 4, and 5 respectively. The second and third elements might be "Thoroughness of treatment" and "Overall quality of care and services", which could have the same or different response options as the first element. If multiple elements in this example question all had the same response options, they could be formatted into a grid, with response options as column headings and the element text as row headings. Such a grid may or may not contain all of the elements for a question. Typically, the fourth element in the example, "When you go for medical care, how often do you see the same doctor?" would be created as a separate question, to promote reusability of questions. If, however, it was thought to be psychometrically important to interpret the patient's response in the context of the list of attributes of a good health care proposed immediately before this element, they could be grouped together as a single question as shown in the example question.

Elements which return a numeric value or a date do not have associated response options. Note that each element produces a single piece of information. If a question asked "Which of the following symptoms does the patient find annoying? (mark all that apply)", and a list of symptoms followed with a check box next to each, then each of the symptoms would represent a separate element, returning a Boolean flag, with one response option corresponding to the Boolean "true". In this case, it is impossible to distinguish between a false and a missing value. The data stored accurately reflects the ambiguity inherent in the question. If the author of the question intended to resolve this ambiguity, at the expense of convenience to the survey-taker, the question could have been structured with each symptom associated with a "yes" and "no" response option, treated as a multiple choice (single answer) element.

The Question Database

The HOI is a non-profit organization working to develop, disseminate, and test public domain tools for measuring the effectiveness of health care. The Henry Ford Health System Center for Clinical Effectiveness has collaborated with HOI to establish the HOI® Outcomes Question Database. The database includes the following information for each question:

1. Question description. A brief description of the question.
2. Question identifier. A three-part numeric code identifying the submitter number, the question number and the element number.
3. Question and response options text and layout;
4. Respondent field indicating the individual who will complete the question;
5. Language in which the question was submitted;
6. HOI acceptance status. This code indicates if HOI has evaluated and recommends this question as compared to others.
7. Subject heading. Subject index terms assigned according to the National Library of Medicine's Medical Subject Headings (MeSH) controlled vocabulary.
8. Origin. Information related to the source of the question to distinguish between the source and the submitter, since they may not be the same.
9. Notes. This field is used to convey potentially useful information concerning the questions development and psychometric properties.

Assignment of Unique Identifiers

It is of obvious importance to a coding standard to assure that each question and each element within questions are assigned a unique identifier. The HOI coding standard uses an identifier with three components: the submitter number, the question

number and the element number. Each component is represented as a non-zero positive integer with no leading zeroes. The whole identifier is represented with the period character as a delimiter. The submitter number and the question number uniquely identifies questions. Within a question, elements are assigned integers in sequence, beginning with 1. Unique submitter numbers are assigned by HOI to each person or institution intending to create or modify questions. These parties can then assign identifiers to new questions using this submitter number. Modifications to pre-existing questions are assigned a new identifier and identifiers are not re-used. The submitters take on responsibility for assuring the uniqueness of codes they assign. They may choose not to actually submit all of the questions they create to the HOI® Outcomes Question Database. This approach permits individuals and institutions to create instruments, assign internationally-unique identifiers for each element on the instrument, and begin use of the instrument without any waiting. This approach also permits easy modification of existing instruments. Forms can be edited by simply inserting or deleting questions in the appropriate location.

Outcomes Data Transmission Standards

Rather than invent a new scheme for outcomes data transmission, we applied the ASTM-E-1238-94/HL7 standard for use with health outcomes data. In general terms, this standard architecture provides a mechanism of sending a linear stream of data which can represent the following set of container relationships: A *message* contains data for one or more *patients*. A patient contains one or more *observation orders*. An observation order contains zero or more *observations* (results). The standard provides field definitions for four "segments" corresponding to these container levels: message header (H); patient (P); observation order (OBR); and observation result (OBX).

IMPLEMENTATION

ASTM Adoption

These coding standards have been incorporated into the ASTM-E-1238-94 data coding and transmission standards approved by the ASTM [10]. The HOI® Health Outcomes Question Database is the only recommended coding system for outcomes data listed in the ASTM-E-1238-94. The ASTM assigned the code "HI" to identify data coded using HOI outcomes data element codes.

Question Database Publication

Access to and use of the HOI® Health Outcomes Question Database has been made available to interested groups, organizations, and individuals. For a modest fee to the Health Outcomes Institute, users can obtain hard copy of the question database documentation. Electronic access to the database will ultimately be offered through the Internet. Users will be notified of the availability of the database through these media.

Incorporation into the AMGA Outcomes Measurement Consortia

One of the early users of the outcomes data coding and transmission standards has been the AMGA Outcomes Measurement Consortium. This consortium includes more than 50 clinics which are actively participating or planning data collection projects for a number of conditions or procedures including total joint replacement, cataract surgery, low-back pain, asthma, and diabetes. To facilitate the adoption of the HOI Question Database/ASTM standards and to expedite the electronic transfer of data for the AMGA data pooling project, the AMGA adopted a policy in 1995 to require data to be submitted in compliance with the HOI/ASTM/HL7 standards. The AMGA outcomes data repository itself was entirely converted to a question-based format, with all results stored using HOI coding standards.

Conversion Software Development

Henry Ford Health System and the AMGA have worked together to develop software tools for standards-compliant data conversion. This data conversion software consists of two main components, both of which are Microsoft Windows-based.

The first software component, the Outcomes Data Conversion Utility (ODCU), is used by individuals or institutions collecting outcomes data to convert their data from ASCII or DBF format into ASTM-E-1238-94/HL7 format. The ODCU includes a question search utility permitting users to search the question database for specific questions by subject, keywords, or identifier. It permits users to define the layout of their input file and map each input field to HOI questions and elements. Once these input formats are defined, data conversion is accomplished by specifying source file, destination file and input format. The software is user-friendly due to the use of the graphical user interface (GUI) and the "wizard" or dialog-style function approach. The

software does not require an attendant while running. The software does not yet support automated modem or Internet-based data transfer.

The second software component, the Standard Outcomes Data Import Utility (SODIU), is used by sites which receive the transmitted ASTM-E-1238-94/HL7 standard data for incorporation into the pooled database. SODIU allows the user to take an ASTM-E-1238-94/HL7 formatted file and convert it into a DBF file containing the HOI Question Database standard codings. Data integrity is verified using an error checking mechanism that allows the user to view error logs and decide to ignore or accept incoming data. Incoming data may be merged with current databases or stored in stand-alone files.

EVALUATION

The Question Database

The HOI Question Database contains a diverse collection of standardized outcomes measurement questions and instruments. As of January 1996, the question database consisted of 1,641 questions. These questions are organized under MeSH subject headings.

The AMGA Outcomes Data Repository

The AMGA Outcomes Data Repository contains pooled data for 7 different clinical conditions. The AMGA uses a standards-compliant question-based architecture which is consistent with the ASTM-E-1238-94/HL7 model. Data has been successfully pooled from 37 institutions measuring outcomes on more than 10,000 patients under the question-based model.

Performance Assessment of Conversion Software

The ODCU and SODIU products were made commercially available for the first time at the July 1996 AMGA Outcomes Measurement Consortium meeting meeting held in San Francisco, CA. At this meeting, the AMGA endorsed this product and recommended it for use by its members. The AMGA leadership anticipates this product will greatly reduce the burden of pooling health outcomes data from their more than 50 member organizations.

Future work will include continued expansion of the Question Database, enhancements to the conversion tools, and providing Internet-based access to the Question Database.

Acknowledgments

This work was supported in part by National Library of Medicine IAIMS Planning Grant #1-G08-LM05720-01.

References

1. Ellwood PM. Shattuck lecture--outcomes management: A technology of patient experience. N Engl J Med 1988;318(23):1549-56.

2. Greenfield S, Kaplan S, Sillman R, et al. The Uses of Outcomes Research for Medical Effectiveness, Quality of Care, and Reimbursement in Type II Diabetes. Diabetes Care 1994;17(Suppl.):32-8.

3. Lahad A, Malter AD, Berg AO, Deyo RA. The Effectiveness of Four Interventions for Prevention of Low Back Pain. JAMA 1994;272:1286-91.

4. NCQA HEDIS Users Group. HEDIS 2.5; Health Plan Employer Data Information Set. Washington, D.C.: NCQA, 1995.

5. Pacific Business Group on Health and William M.Mercer I. Employee Medical Plan Satisfaction Survey. San Francisco, CA. Pacific Business Group on Health, 1991.

6. O'Connor GT, Plume SK, Olmstead EM, et al. A Regional Intervention to Improve the Hospital Mortality Associated with Coronary Artery Bypass Graft Surgery. JAMA 1996;275(11):841-6.

7. Buckstein D, Luskin AT. Successful Treatment of Chronic Asthma Using Point-of-Service Outcomes Analysis. Journal of Clinical Outcomes Management 1995;2(2):45-60.

8. American Health Consultants. Patient interviewers find out what others think about one practice. Patient Satisfaction & Outcomes Measurement in Physician Practices 1996;2(1):1-12.

9. Davies AR, Ware JE. GHAA Consumer Satisfaction Survey. Second ed. Washington, D.C.: Group Health Association of America; 1991;

10. ASTM. Annual Book of ASTM Standards. Standard Specification for Transferring Clinical Observations Between Independent Computer Systems E-1238. Philadelphia: ASTM; 1993.